Image Cover Sheet

CLASSIFICATION	SYSTEM NUMBER 510918
UNCLASSIFIED	

TITLE

COGNITIVE TASK ANALYSIS OF THE HALIFAX-CLASS OPERATIONS ROOM OFFICER
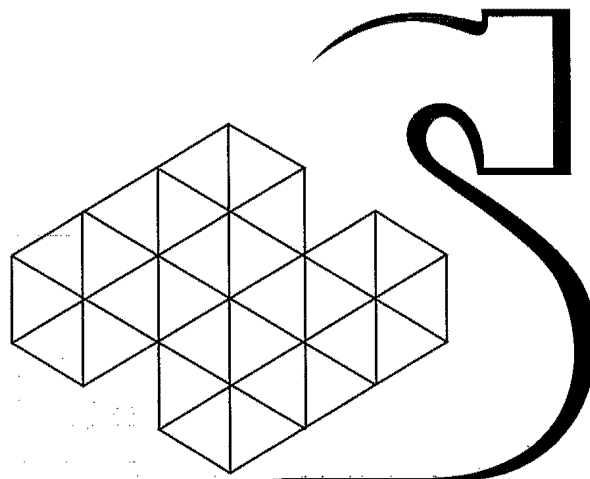
System Number:

Patron Number:

Requester:

Notes:

		Form Approved
Report Documentation Page		OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **10 MAR 1999**	2. REPORT TYPE	3. DATES COVERED **00-00-1999 to 00-00-1999**
4. TITLE AND SUBTITLE **Cognitive Task Analysis of the HALIFAX-Class Operations Room Officer**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **HumanSystems Incorporated,111 Farquhar Street,Guelph, Ontario, Canada N1H 3N4 ,**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**		
13. SUPPLEMENTARY NOTES		
14. ABSTRACT		
15. SUBJECT TERMS		

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**49**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

HUMAN*SYSTEMS*

Incorporated

Humansystems

111 Farquhar Street

Guelph, Ontario

N1H 3N4

Tel: (519) 836-5911

Fax: (519) 836-1722

http://www.humansys.com

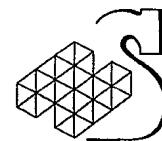## Cognitive Task Analysis of the HALIFAX-Class Operations Room Officer:

PWGSC Contract No. W7711-7-7404/001/SV
Order No. 7404-03, 7404-06
PWGSC Contract No. W7711-8-7433/001/SRV

March 1999

COGNITIVE TASK ANALYSIS OF THE HALIFAX-CLASS
OPERATIONS ROOM OFFICER

by:

Michael L. Matthews, Robert D.G. Webb, and David J. Bryant

Human*systems*, Incorporated
111 Farquhar St., 2nd floor
Guelph, ON N1H 3N4
(519) 836-5911
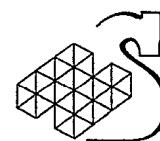
Table of Contents

Executive Summary

This report details the results of a Cognitive Analysis of the Operations Room Officers (ORO) position in the HALIFAX class frigate. The purpose of the analysis was to better understand cognitive aspects of Command and Control (C2) in the Operations Room (OR) as a prelude to preparing evaluation measures and methods to assess future Command and Control support systems. The ORO position was selected for analysis from among other OR positions as being the position with the most comprehensive responsibilities, most likely to provide a broad insight into C2 functions and to generalise to other positions. This choice was based in part on an earlier DCIEM report on opportunities for improving C2 support for the OR identified by members of the OR Command team (the Commanding Officer, ORO, Surface Weapons Controller, and Anti Submarine Weapons Controller).
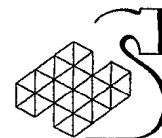
This cognitive analysis was based on interviews with four pairs of experienced OROs, with each interview session lasting two days. Interviewees were walked through an operational scenario and prompted to describe their cognitive activities based on their experience with similar events to those in the scenario. The scenario was derived from a training exercise at the HALIFAX land based training simulator and stepped a Canadian Task Group (TG) through a series of threat responses as the TG passed through coastal waters with a high level of civil air and surface traffic. Threats included sub-surface, surface, and air in single and multiple combinations with periods of intervening surveillance. The ORO position in question was in a HALIFAX class vessel in the forward screen, with no TG warfare duties.

Although, necessarily, conducted in the context of one position, and constrained by current technology and procedures, the objective was to identify functional goals for the ORO position and associated cognitive needs. Based on a previous review of the literature for C2 systems, four key functional areas were used to focus the interviews and for development of evaluation measures. These were Situation Awareness (SA) for any pertinent information, all forms of Communication (CM), all levels of Decision Making (DM) and Mental Workload (MW). These terms are described in the report.

Interview data were organized into data tables (shown in an Annex to the report) by identifying first an ORO's generic goals (for a particular aspect of the scenario) and then the related needs for awareness, communication, and major decisions for achieving that goal. Preliminary measures and methods to assess the degree to which these needs are supported are then given. Methods of data capture are based on what should be feasible in a dedicated test bed or training simulation context.

Selected observations arising from the interviews (and expanded in the report) include the following:

- OROs employ a variety of mental pictures and models to achieve their cognitive goals. These mental representations vary in terms of level of detail, domain, and abstraction and may be modulated by career path, experience, other team members, mission planning, watch hand-over, and recent events.
- Mission preparation has great significance for establishing mission related mental models.
- Updating Situation Awareness when coming on watch is critical for updating the ORO's mental models.
- Information systems need to provide less data and more "information" i.e. pertinent data meaningfully integrated in terms of the users mental model(s). This implies a need for functional support that can be adapted by "educated" users to individual, contextual and mission related needs.

- For Situation Awareness, major functional requirements (currently not well supported) include information acquisition and integration, cognitive fit of the available data to the mental model(s) of the user, regaining awareness after switches in attention between different areas of focus, and alerts for significant changes in unattended areas.
- Background tasks such as maintaining situation awareness of the evolving operation or dealing with incoming text messages, can be differentiated from foreground or threat-related tasks. These tasks interact and require the ORO frequently to switch attention and change mental models. OROs need to be able to switch between foreground and background tasks with seamless integration of data.
- A major ORO function is to manage the overall OR team threat response rather than to be directly involved in the details of responding to particular threats (although this requirement may change for multiple threats).
- Common implicit intent and understanding among OR team members has particular significance for communication effectiveness.

Many of these points probably apply, to some degree, to other OR positions.

The next recommended step will be to verify the present findings and proposed measures with one or more experienced naval SME's. This would be followed by a brief, practical, trial of the measures and methods outlined in this report, using a small, selected segment of the present scenario. There would be three purposes for this trial. First, as a general verification of the measurement approach and to further refine the measures. Second, to check that data can be captured using simulation resources available, first in the Halifax land based trainer and then, perhaps, on board ship (in harbour or at sea). Third, to work through the analysis of a mock data set to illustrate how baseline data might be derived for the present system.

The literature on human reasoning appears to offer rich insights into the challenges faced by OR members and yet to remain largely untapped by naval decision making research. Therefore as a parallel step we recommend a review of the psychological literature on mental models and other forms of mental representations in the context of the management of advanced multi-task systems. Areas of this literature have evolved somewhat in isolation one from the other and from application domains. This review would contribute to issues related to both evaluation methodology, as well as research and design implications in the context of the OR.
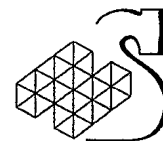
1. Scope and Background

This work was commissioned by the Defence and Civil Institute of Environmental Medicine (DCIEM) as part of a standing offer contract to support ongoing Defence Research and Development Branch efforts to improve the application of human factors engineering in DND CCIS projects (Contract No. W7711-7-7404/001/SV). The Technical Authority is Carol McCann (DCIEM).

1.1. Background

The Navy has identified decision support in the Operations Room (OR) as a focus of attention for the upgrade to the HALIFAX Class Canadian Patrol Frigate (CPF) anticipated for the 2005-2010 time frame. However, there is a lack of knowledge concerning the information which forms the basis for OR decision making, the sources of this information, its criticality, and the strategies used to gather this information and make decisions during various phases of operations. Consequently, the requirements for a future Command and Control Information System (CCIS) that will enhance decision support cannot be clearly determined at the present time. Further, the measures by which any decision support tool or technique would be assessed need to be established. As a first step to rectifying this deficiency, an analysis of OR decision tasks was carried out. The initial study was limited to the ORO functions (and position), as being the most critical and complex. The analysis may be extended to other positions in the future. This work was carried out under contract for the Defence and Civil Institute of Environmental Medicine (DCIEM).

1.2. Overall Objectives

The purpose of this part of the project was to undertake a cognitive task analysis (CTA) of the ORO function in the HALIFAX class. This CTA is intended to provide a basis for the development of detailed measures and methods for assessing the performance and effectiveness of the CCIS that supports ORO functions. (It is likely that the CTA will provide design insight into the functionality required of the next generation of system.) The focus of the analysis is based on four primary performance factors identified in previous work: Situation Awareness (SA), Decision Making (DM), Communication (CM) , and Mental Workload (MW). The overall framework for this analysis was provided by mission scenarios.

1.3. Report Structure

This report provides an overview of the methods used, key concepts, data reduction strategy and an analysis of the data obtained. Where possible, and to maintain clarity, cross references are provided to related reports which provide more detailed information in selected areas, such as the scenarios used during the CTA.

This report is supplemented by Annexes detailing the outcome of the CTA, in the format described in this report.
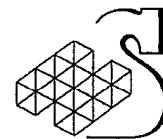
1.4. Limitations

This project is limited to the Operations Room Officer (ORO). Other positions within the OR and operations from other locations within the ship are excluded. The scenario covers only ongoing surveillance and responses to single or dual threats. Mission planning and preparation responsibilities are not included.

1.5. Review of work accomplished to date

Two previous projects and their associated reports should be read in conjunction with this report. The first report (Reference A) described the outcome of interviews with a cross section of approximately 18 OR team members comprising what is sometimes called the "Back Row" team. This team comprises the Commanding Officer (CO) of the ship, the Operations Room Officer (ORO), the Surface Warfare Controller (SWC) and Anti Submarine Warfare Controller (ASWC). That report describes the general responsibilities of each of these positions in relation to other OR team positions, concerns with the functionality of the current system, and opportunities for improvement. The interviews were based on a scenario similar to the one used for the current CTA. Major differences were that the current scenario omitted any mission planning, placed the interviewees ship in a Task Group (TG) context rather than as an isolated ship, and went into greater detail for each scenario event. The second report (Reference B) details the scenario and data capture methods to be used in the present CTA.
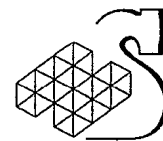
2. Referenced Documents

A. Webb R.D.G., McLean D.N., McCann C.A.,
"Deficiencies in Command and Control support to the Command team in the HALIFAX class."
Report to DCIEM-PWGSC Contract No. W7711-6-786/01- SRV, November 1997.
Humansystems Incorporated

B. Webb, R.D.G, McLean D.N., Matthews, M.L., McCann C.A.
"Halifax Class Upgrade 2005, Cognitive Task Analysis of the ORO Position: Outline Method".
Report to DCIEM-PWGSC Contract No. W7711-7-7404/001/SV. July 1998.
Humansystems Incorporated

C. Matthews M.L., Webb R.D.G., McCann C.A.
"A Framework for Evaluating Military Command and Control Systems",
Report to DCIEM-PWGSC Contract No. W7711-5-7253, June 1997.
Humansystems Incorporated

D. Endsley M.R. 1995.
"Towards a Theory of Situation Awareness in Dynamic Situations."
Human Factors, 37 (1), 32-64.

E. Crandall B., Klein G., Militello L.G., Wolf S.P. 1994,
"Tools for applied cognitive task analysis."
Klein Associates: Fairborn, Ohio.
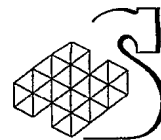
3. Technical terms and abbreviations

Abbreviations

C2:   Command and Control
CTA: Cognitive task analysis
MOE: Measure of effectiveness
MOP: Measure of performance
OR: Operations room
ORO: Operations room officer
OSD: Operational sequence diagrams
ROE: Rules of engagement
SAGAT: Situation awareness general assessment technique
SME: Subject matter expert

Technical terms

Warfare Area: a major area of surveillance and operational responsibility. Normally divided into air, surface or sub-surface. The term is used both to describe the external focus of activity and the internal resources (personnel and equipment) that support operations in the area.

Warfare director: the senior officer on watch in the OR responsible for a warfare area for the ship. May be an officer or a petty officer. There will also be a TG warfare director. This person may be located in the TG command ship, or the role of warfare director for one or more warfare areas may be assigned to different ships in the TG. For the purpose of this study, all warfare director responsibilities were assigned to the TG ship.
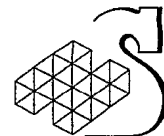
4. Method

4.1. Goals of CTA approach

In this section we briefly outline how the CTA methodology provides a basis for understanding and decomposition of major ORO tasks. This in turn results in the development of specific MOP's that provide indicators of OR effectiveness. As future systems evolve, the key elements of this decomposition and analysis can be used to guide the development of user requirements, provide design feedback throughout the development cycle and, subsequently, to evaluate the effectiveness of design, staffing or procedural changes on OR performance.

4.1.1. Measurement areas for C2 performance

In a previous review of the broad literature on MOE and MOP for military C2 systems, we concluded that there are four key areas for user performance and hence system effectiveness (Reference C). These are Situation Awareness (SA), Communication (CM), Decision Making (DM) and Mental Workload (MW) and each merits measurement. A brief description of each of these follows; with more detailed descriptions in References C and D. As will be noted, these areas of performance are inter-related. Improvements in one area may well result in changes in the indices for others. Therefore, measures should not be regarded as isolated one from the other.

_Situation awareness_: the perception and extraction of pertinent information from the operating environment to support present and future decisions. SA also encompasses the integration of the extracted information into a picture (literal or conceptual) that is comprehensible to the individual and the projection of that picture forward over time. Originally SA was conceived in terms of the relatively short time frames (seconds and minutes) and spatial information critical for air combat. More recently, the literature has seen the concept broadened to encompass much longer time frames (hours, days and even weeks) and a wider range of information (equipment and personnel status, and any other problem-related information). In this project, the broader concept is used.

Typical failures in situation awareness that affect performance include:
- not seeking the appropriate information
- not detecting critical information
- not recognising the criticality of information
- attending to the wrong information
- not integrating all appropriate information into the current understanding of a situation
- failing to appreciate how the situation may evolve in the future
- failing to anticipate future information needs.

Communication: the means by which the ORO acquires or transmits information. Communication may be conducted verbally (directly face to face or over networks), in writing or graphically (paper and electronic) and by direct facial expression and gesture. This involves OR personnel, external communications to the task group, interaction with the CCS, or any other source of information in terms of state-boards, manuals, or displays. Communication may be described as the process of acquiring information to gain and maintain SA as well as a means of giving or receiving direction. Even if the end result for SA may remain the same, communication variables may affect individual or team resources that must be used to gain that awareness. Functionally, the communication process may vary from system to system depending on the technology and design philosophy
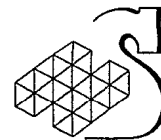
employed. Measures chosen should permit comparison of user performance in systems employing different technologies to enable users to acquire the same information. Thus in one system, users may communicate by semaphore, in another by voice, in another by "post-it" notes and another by symbols on a CCS screen. Measurement of speed, accuracy and timeliness of information acquisition for different systems may be compared. Measurement overlap exists between SA and communication in the area of display configuration. For example, one display format may provide better cognitive fit with the mental model employed by the user to integrate pertinent information to solve the problem(s) with which he or she is presented. Such a display format could very well affect measures of SA (faster and more complete detection of pertinent information, better comprehension) as well as communication (faster detection, faster comprehension as a result of fewer queries, lower potential for misunderstanding). Major problems in communication are:

- errors of omission
- transcription errors
- incorrect information transmitted
- delays in required information
- rapidly developing backlog of information

Decision making: The scale of "decision(s)"considered can vary considerably. For this project, the term "decision" is taken to mean a choice made by the ORO. Such choices may vary in scale such as the choice to range in or range out, to leave the SWC to deal with a particular threat, to double check a report or not. A more telling decision may be to fire or not. The outcome of some decisions may be visible and readily apparent, e.g. when the ORO takes specific actions or gives commands to effect changes in the operating environment (changing course, launching a weapon). Other decision outcomes may be less apparent, e.g. when the ORO assesses a threat, or selects a specific piece of information from the environment but, nevertheless, affect the threat response. The view taken in this project is that all critical choices made by the ORO in the pursuit of preparation for or actual threat responses merit consideration for measurement.

Major problems with decision making may include:

- decision is based upon incorrect information
- decision fails to factor in relevant critical information
- decision is too slow
- decision is inappropriate
- decision is biased.

Mental Workload: This represents the ability of the operator to achieve goals and perform tasks in the time available and within their personal information processing bandwidth. By this we mean either generic human bandwidth (such as the rule of thumb 7 plus or minus 2 items of information) or individual differences in bandwidth (e.g. novice vs. expert), or situationally determined bandwidth (e.g. narrowed attention under stress). An excess of tasks, or task complexity, over the time available to do them increases time pressure. As a result tasks may not be properly completed, delayed, or may have to be time-shared with other competing tasks. A mark of the expert is not only increased bandwidth but better ability to re-organise tasks (drop, add, re-assign) to achieve goal priorities, provided the system permits. These elements interact. For example, an expert may organize information differently and be less susceptible to stress effects. Complex, time consuming or error prone communications will not only affect SA but also increase workload. Prolonged periods of high time pressure and resulting high workload result in:

- attention narrowing
- a reduction in or complete loss of situation awareness
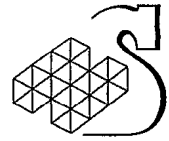
- hasty and incorrect decisions

At the other extreme, extended periods of low workload typically result in loss of vigilance. Hence the goal of system design is to create a task environment with moderate workload levels and manageable periods of high workload peaks. A major role of the ORO in this respect is to manage not only his own workload (perhaps by shedding tasks according to some scale of priority), but more important, to manage workload among the OR team by re-assigning work as appropriate.

For either extreme, high or low, measurement of mental workload is an important diagnostic indicator for the existence of problems in the other areas. Conversely, improvements in SA, Comms, or DM should result in mental workload falling within the desirable range.

4.1.2. Relationship to core Navy C2 concepts

Our literature search and meetings with SMEs has uncovered three widely used terms in the discussion of C2 support systems for navy operations – *situation and threat assessment (STA)*, *resource management* and *data fusion*. This section outlines some potential relationships between these concepts and the four core measurement areas identified above.

While S*ituation and Threat Assessment* encompass elements of situation awareness they also implies a level of information analysis and interpretation which goes beyond the usual definition of SA. This further analysis involves a conscious evaluation of the data made available by SA in terms of its tactical or strategic importance. As a result, tactical context is seen to influence the way in which *data* become transformed to *information*. For example, mental picture(s) (see below) generated by the ORO from data provided by three surface radar returns in formation may be quite different depending upon other relevant tactical intelligence. For example, in an area known to be populated by fishing boats, radar contacts may be interpreted as uninteresting "noise" and discounted. In an area of hostilities, in which enemy surface units have been reported by other intelligence, the same data will likely be highly prominent and salient in the ORO's mental picture.

Chalmers (DREV briefing to HSI) outlines Situation Interpretation, Situation Projection, Situation Monitoring and Diagnosis as major processes in Situation and Threat Assessment. Of these, *situation monitoring* seems to correspond most closely to Endsley's level 1 SA (Reference D), *situation interpretation* to SA level 2 and *situation projection* to SA level 3. The *diagnosis* component seems to have no direct counterpart, but can be considered a special case of decision making for the purposes of measurement. However, the ambiguity inherent in any one formulation of the cognitive relationship between acquiring awareness and making decisions is underscored by comparisons such as this. The level at which a "decision" is seen to exist seems important here. For example, is the choice of which information to seek (and where to seek it) as one acquires awareness of the "situation" in itself a "decision"?

Resource Management (RM) is a concept relating to the management of both equipment and personnel in the execution and implementation of operational actions. Component elements of RM include: engageability calculations, target-weapon pairing, and assignment of resources and fire control. As such, all of these processes include essential elements of SA, CM and DM; hence the MOPs prescribed below for these elements will also be key indicators of processes that underlie successful RM.

Data fusion can occur at four levels: object refinement, situation refinement, threat refinement and process refinement. CTA can be expected to assist designers of data fusion refinements to understand the cognitive processes of users appropriate to the differing levels of data fusion, and by
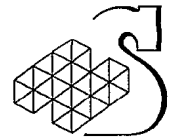
cataloguing relevant information requirements to support each level of data fusion. The applicability of CTA will probably be greater in the area of process refinement.

At present no specific measures or methods have been developed for the evaluation of situation and threat assessment. We believe that the most promising approach lies with the freeze probe methodology and SME review (see below). Many aspects of situation and threat assessment can also be evaluated by measuring the related processes of SA and DM.

4.1.3. How CTA provides the necessary information for the key MOPs.

The major steps in the CTA involve:

- acquisition of domain knowledge by the CTA team

- development of representative scenarios (context and activity sequence) based upon mission analysis plus observation and interviews with experienced OROs.

- stepping ORO interviewees through the sequences of events that comprise the scenarios

- identifying ORO goals and related decisions and tasks.

- identifying the information needs to support critical decisions and tasks.

- identifying critical incidents, errors or other problems for ORO's based on interviewees' operational, exercise, or simulation experience when pursuing such goals

- identifying appropriate measures of the ability of the system to support users to:
  - acquire the required information at the appropriate time and in the appropriate format;
  - interact effectively with the sources of the required information;
  - make appropriate and timely decisions;
  - accomplish identified goals
  -- maintain workload within a desired range.

A second goal of the CTA was to identify some future research priorities in Halifax class OR C2 operations. Through the CTA we have been able to identify a number factors that affect ORO and, by implication, OR performance, and which will require greater analysis and/or understanding in order to provide future system design guidance and possibly to identify appropriate MOPs. Such issues are aspects of the OR/ORO operation for which little or no research is available currently.

4.2. Summary of method

4.2.1. Scenario approach

In outline, the proposed method follows the eclectic approach of Crandall et al (1994) (Reference E), who advocate structured interviews based on a standardised scenario, with probe questions to elicit different information or cognitive requirements. The chosen scenario stepped ORO interviewees through a series of threat events for a HALIFAX class vessel operating in a Task Group in a hostile littoral setting. This setting has day-to-day activity (commercial air lane, fishing vessels, oil platforms with helicopter traffic) that must be monitored, and from which will emerge friendly, neutral, and hostile contacts (described fully in Reference B). Interviewees were asked to identify their information needs in terms of key variables underlying Command and Control: Situation Awareness, Communication, Decision-Making and Mental Workload. They were also asked to identify the goals they had in mind, critical incidents, errors they had experienced or seen other ORO's produce, particular difficulties they faced in trying to achieve a goal and expert-novice differences, and workload, where excessive.
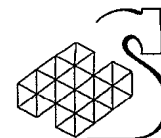
This standardised methodology permitted comparison between the answers of different interviewees. The fidelity of the scenario to training or other experiences helped interviewees visualise the tasks in question. The probe questions were designed to reveal cognitive issues relating to the four key variables. To these ends, the details provided in the scenario were designed to promote and guide discussion rather than to be entirely realistic.

Eight OROs were used as SME interviewees, four from the West Coast and four from the East Coast. SMEs had at least six months experience in a HALIFAX class vessel since completing their qualifying course. SMEs were interviewed in pairs partly to economise on time but more to promote discussion of different perspectives and thereby enhance the validity of the data obtained.

4.2.2. Schedule of Interviews and Setting

The interview process required two days with each SME pair. With ten scenario events and allowing two hours for familiarisation with the CTA process, this allowed approximately one hour for each event in the scenario. Settings were a quiet office environment with a black board and table space. We provided a simple mock-up of a CCS workstation, a layout diagram of the OR and pertinent charts of the area in which the scenario was set. Each participant had a copy of the scenario activity sequence and the main headings under which insights were desired.

Interviews comprised the following parts:

- **Initial Briefing:** This outlined the purpose and structure of the sessions and the scenario and explained that the purpose of the scenario was to provide a review framework and not to test participants' knowledge or reactions. Key terms such as Situation Awareness were explained. Each participant was asked to complete a data sheet outlining pertinent professional experience.

- **Knowledge Review:** Using a schematic of the OR as a guide, SMEs were prompted to describe the various types and sources of information needed by an ORO.

- **Detailed Discussion of each Event.** This formed the majority of the interview time. Working through events one by one, we tried to establish aspects such as:
  - Key decision points and cognitive goals.
  - Information required at those decision points
  - Critical cues or patterns and trigger points
  - Communication needed to acquire key information or make decisions
  - Tasks requiring expert judgement and decision support rather than routine skills.
  - The sequence of knowledge acquisition.
  - Expert-novice differences.
  - Mental operations to retrieve, store, transpose, integrate, or model information.
  - Meta-cognitive processes for directing mental effort and attention.
  - Users mental models and methods of problem visualisation.
  - Error patterns, near misses, stress points, and critical incidents.
  - Situations which require multi-tasking and time-sharing.
  - Information which must be memorised.
  - Monitoring activities for optimum team functioning.
  - Planning next actions for self and team

These data was captured using a data sheet in which the columns formed the major areas of interest for data capture and the rows contained the successive tasks and activities that were sequentially triggered by the events of the scenario. The layout of a data table is shown below (Table 2).
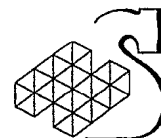
The briefing package provided for each interview can be found in Reference B

4.2.3. Interview details

Interviews were conducted in separate visits, first to MARLANT and then MARPAC by the same team of interviewers. OROs were interviewed in pairs over a period of two days. Two sets of interviews were conducted with different pairs in each location, which yielded a total of eight participants. OROs were all enthusiastic and engaged participants, who took time to think through their experiences before responding.

The paired SME format was successful in generating active discussion on the issues involved and enabled different perspectives and operational experiences to be brought to bear, thereby enriching and giving higher validity to the data obtained. In some cases, where there was dissent, this was noted. In other cases, we tended to rely more on the data obtained from the more experienced member of the pair. Participants reviewed each activity sequence prior to starting detailed discussion about the cognitive issues and, generally, agreed that scenario activity sequences were representative of the way actual events occurred. Minor changes were made to procedural sequences, or to combine activities presented as separate when, in practice, these activities might overlap or follow so quickly one after another as to be virtually indistinguishable.

At the end of the interview, when the OROs had a good understanding of the process and what we were trying to achieve, they were encouraged to add any additional comments or general observations relevant to the CTA goals.

Because it took longer to step through all of the scenario contact episodes than was anticipated, it became necessary to be more selective of scenario segments in later interviews, to ensure we obtained the broadest data sample, given the limited opportunity for access to SMEs. It also became apparent during the first interview that OROs separated background ongoing tasks (such as dealing with incoming text message traffic) from foreground threat related tasks (such as calculating missile release). Furthermore, OROs made reference to earlier activities as providing them with their mental models for situation assessment and response planning. These earlier activities including mission planning and preparation, earlier experiences during the same mission, and information acquired when preparing to take over the watch. Because mission planning and rehearsal were explicitly excluded from the terms of reference for this project, related issues were not addressed during the interviews, despite their apparent significance. However, the information gathering in preparation for coming on watch and the distinction between foreground and background tasks were examined with the interviewees. The table below shows the major segments of the scenario that were covered in each set of interviews.

Location / pair	Scenario component			
	Coming on watch	Background tasks	Single threat	Multiple threat
East Coast A	YES	YES	YES	
East Coast B			YES	YES
West Coast A	YES	YES	YES	YES
West Coast B			YES	YES

Table 1: Scenario segments reviewed with each interview pair
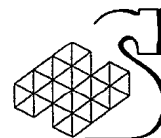
4.2.4. Framework for data reduction

Interviewee responses were recorded individually by the two interviewers directly into data tables (see Table 2) prepared for each scenario event. (That is with the sequence of activities comprising the event listed down the left-hand side of the data table.) On the same day the data were collected, the two interviewers reviewed and reconciled the two sets of data with each other. In general there was a high level of consistency. In cases of disparity, questions were noted and addressed during subsequent interviews. Other than some local differences in individual operational experiences, there were no marked differences that we could observe between the data obtained from MARLANT (East Coast) and MARPAC (West Coast).

| Scenario Event	Goal(s)	Awareness needs	Decisions	Communication	Mental Workload	Comments
Activity #1	Interview responses to probe questions about activity					
Activity #2						

Table 2: Outline interview data table

The next task of the analysis was to categorise the data for each of the activities comprising each scenario event. This was done by aggregating the data from the eight interviews and assigning it to the appropriate category (columns for goals, awareness needs, decisions, etc).

At the end of this process, we had achieved integration of all data collected and produced a number of tables of collated raw interview data (not shown) for each threat event as well as goals for background tasks and coming on watch. Reviewing these data, it became clear that superficially different external event triggers were driving the analysis and that this resulted in a lot of descriptive repetition. In other words, events that might differ by warfare domain triggered identical cognitive responses because OROs had similar goals across different warfare domains. For example, monitoring the progress of threat identification within the OR occurs for surface, sub-surface air and also multiple threats. In each case the ORO has the same goals in mind, but the focus of attention (in terms of information sources and other members of the OR team) and timelines differ. It thus became clear that there was a need to bring together generic behaviours and to organise these by *ORO goals* rather than tasks in response to specific external triggers. By doing so, we hope that a simpler and more organised picture would emerge of the ORO's C2 cognitive needs and that these would generalise in functional terms to a variety of existing and possible future C2 tasks. See sample data tables in Annex A.

The goal driven aggregation of the data that has emerged represents a preliminary and initial attempt to bring together the key elements of the CTA. Refinement and further development of this analysis will be required in the next stages of the work.

4.2.5. Goal analysis

The initial key concept which drives the analysis, is the assumption that the important ORO C2 behaviours, information processing and actions are *goal directed*, that is tasks are performed as part of a series of steps in achieving a goal. Hence the first step of the analysis is to determine the goals and then to look at how the goals are accomplished, what information is required to support meeting the goal and what critical decisions are made.
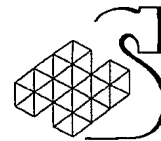
The initial challenge in such a goal-means decomposition is to determine an appropriate level for the goal analysis. Some compromise and judgement needed to be made between selecting too high a goal (e.g. "*achieve mission objective and return safely to harbour*") and too low a goal (e.g. "*direct attention to the CCS screen*"). We tried to identify intermediate, generic goals that serve the achievement of broader goals.

One problem with any approach to decomposing a complex mission into component goals is the risk of successive recursion down to the most minimal of goals. We tried to avoid this by stopping the analysis at the level of decomposition that appears to us to make the most sense from the point of view of performance measurement. In particular we took into account what OROs defined as major goals and important areas by which the system performance could be evaluated. Further we focussed on analysing intermediate-level goals which serve as wide a variety of mission related tasks as possible. An example of such a goal is "*manage the OR team*". Such activities occur in almost all operational contexts and has certain stable characteristics which always apply.

The major alternative approaches to a goal-driven analysis would have been to use either a more detailed task-flow analysis, or a more abstract critical decision analysis.

In the present case, a task flow analysis would be constrained and ordered by the specific sequence of events that comprise our mission scenario. The task flow analysis has an advantage in that it describes fairly completely the more or less sequential activities that underlie each externally triggered event. As such, it provides a comprehensible road map of the way tasks are linked and typically ordered in an operational context. One major disadvantage is that many tasks are more or less repeated in a variety of contexts, which leads to a high level of redundancy in the data presented and reduces the focus on critical behaviours. Further, from the perspective of identifying issues that are critical for future systems design, it makes the task of extracting generic design issues much more difficult if the analysis is task based, rather than goals-based.

An analysis based on critical decisions only was also considered inappropriate for the present context. The major reason is that it provides too restricted a focus compared with the broad range of issues that are uncovered by a more generic analysis centred on goals. There remains the possibility that if some critical decisions have not been analysed in sufficient depth with this initial CTA, subsequent analysis using the critical decision approach is not precluded.

Future system improvements will need to be geared initially to facilitating the achievement of operator goals, rather than the current specific tasks that comprise these goals, for two reasons. First, not all tasks can be anticipated in advance. This is particularly true in the context of emerging technologies that may transform existing detailed operational procedures, job descriptions and staffing structures. Second, the nature of the ORO's general work domain needs a great deal of flexibility in the way problems are defined, prioritised and to some degree executed. Thus, an analysis at the task sequence level runs the risk of prescribing the work of the ORO, whereas what needs to be captured is information that will allow the ORO to be flexible in response and allow for creativity and problem solving in dealing with unforeseen situations. Note that interviewee responses tended to be couched in terms of existing technology and procedures rather than fundamental functional terms of information and decision needs. Therefore, one of our major initial goals has been to make a preliminary identification of the generic underlying functional needs of OROs that impact upon system performance. In this respect, this process of extraction and abstraction may have common elements with *cognitive work analysis* and *ecological interface design*.
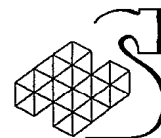
5. Results

This section comprises the following components:

- Discussion of "mental representations or models" used by the ORO to support situation awareness and decision making.
- Explanation of data tables.

5.1. Mental representations

Before proceeding with the description of the CTA data tables, we believe it is necessary to review an important cognitive construct that emerged from the CTA, that is the OROs' widespread use of various forms of mental representations. Frequent allusions were made to "mental pictures" or "mental plans" by OROs to support their situation awareness and their attainment of short and long term mission related goals. In this section we provide an overview of this issue now so that the reader will have a greater appreciation of the measurement issues that follow.

The structure of this section is to initially provide an outline of the context for exploring this issue; this is followed by a section that provides some working definitions around the concept of mental representations. We then provide a preliminary categorisation and description of the various forms of mental representations that OROs may use. Finally, we outline some research issues relating to how future system designs might better accommodate the ORO in the acquisition and maintenance of internal representations of selected features of their current mission environment or projected task sequence. A first step in this research would be to review the literature on mental models and spatial visualisation in their support of selected aspects of human reasoning and decision making.

5.1.1. Operational context

The basis for highlighting this issue was the frequent reference by OROs to their use of some sort of internal representation or mental model of selected aspects of their total work environment, in order to facilitate decision making and/or the management or execution of the task at hand. Such internal representations occurred for all ORO goals and tasks, not just those related to threat assessment. (For example, reviewing incoming text messages or anticipating the personnel resource allocations for the watch.) While the nature of the internal representation was never expressed with absolute clarity across different ORO interviewees, certain consistent themes emerged to suggest that the representation is not one single, simple construct but takes a variety of forms that depend on circumstances and the ORO's goals. For convenience, we have categorised these representations along three dimensions or continua, which may vary in their permutations or combinations for any particular representation. These dimensions are: *'level of abstraction'*: (i.e. a concrete image vs an symbolised model vs an abstract schema), *'level of focus'* (i.e. "ranged" in or out, the ship in isolation vs theTG as a whole, immediate time frame vs future) and overlapping *'domains of interest'* (i.e. warfare and resource areas). In the following sections, we elaborate upon these aspects of such representation by providing several examples drawn from the CTA.

The need for these multiple forms of mental representation of the operating environment appears to be driven by both strategic and tactical requirements that demand that the ORO pay attention to the following:
- both local and more distant (space and time) events
- integrating information about events across the air, surface and sub-surface domains
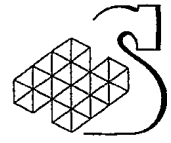
- taking a ship or task group perspective on the former, perhaps depending on warfare responsibilities assigned to the ship
- managing foreground and background tasks, OR work flow and operation
- managing resource issues in terms of personnel, sensors and weapons.

The CTA analysis clearly revealed that the ORO is constrained in the maintenance of such multiple perspectives by existing information systems that provide, at any one time, a selected, limited, scaled view of the external environment. That is, existing systems provide little to support the ORO in the creation and maintenance of either integrated internal representations across domains of interest or across time. They also do not provide the necessary cognitive compatibility to permit OROs to quickly switch from perspective to another.

A further, consistent theme that emerged from the analysis was that OROs use such internal representations on a continuing basis, switch back and forth between them frequently, and struggle to maintain their fidelity and currency. Given the centrality of such representations to the OROs cognitive working domain, we believe it necessary to describe these in more detail and try to categorise them in a manner that will be useful for the development of measures, for framing important research questions and for informing future system design. To facilitate a meaningful categorisation, we provide in the next section some definitions and terminology that we hope will provide some clarity and insight into what can be a potentially complex issue.

5.1.2. Description and definition

Working distinctions have been made in the literature between a *'mental picture'*, a *'mental model'*, and a *'schema'*, which differ among themselves in terms of their level of abstraction/generality and level of 'conceptualness' vs. 'perceptualness'.

A *'mental picture'* is the least conceptual and most specific. It is exactly what the name implies, a graphic visualisation of some aspect of the existing operational environment, or projection of a future state, or recall of an earlier state. The mental picture contains perceptual information but little if any conceptual information, it can only represent what things look like and it has a particular viewpoint. Thus, it is a simulation of a specific view of a specific system at a specific time only, although a mental picture can have some dynamic properties and represent movement. More *conceptual* forms of representations are required to understand what things mean.

The picture may be simple or complex, directly reflecting the data on which it is built (e.g. an image of the raw radar return) or constructed (visualising the data return of a contact as a vessel at some distance and bearing from the ship). The picture will also be dependent, in large part, on the format and level of fragmentation with which the information to be integrated is presented. It is also subject to the usual cognitive principles of selective attention, limited capacity and gradual decay in memory.

A *'mental model'* has been defined as an internal representation that has a similar relationship-structure to that of the processes it imitates. Such models are thought to underlie many aspects of human reasoning and are seen as paralleling the working of the processes they represent. Since the underlying processes may vary considerably from the simple to the complex, concrete to the abstract, the mental model itself will vary accordingly. A mental model is a conceptual representation of a specific system or situation; it contains information about the elements (objects, people) that are actually in the system, it represents relationships that really exist between the elements. The general purpose of the model is to represent the physical nature of the system (things, actions, etc) and the conceptual structure (how are things related to one another). Like a
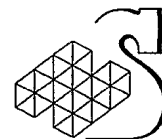
physical model, the mental model does not imply any specific view or interpretation of the system; rather it can be potentially viewed from any perspective. For example, the ORO's mental model of the OR would contain tokens for the people and things in the room (SWC, ASWC etc, the hardware and physical systems), and represent the relationships between them. The mental model would also represent the processes that take place in the OR and how all of the components in the OR interact to make it work.

The generation of such models will be partly derived from the creativity of the individual ORO, and his or her previous meta-cognitive strategies (intuitive or overt) in the management of information complexities while solving problems. This will be modulated by training and experience and be supported, or constrained, by the underlying information system capabilities.

A 'mental schema' represents the most general and abstract form of the internal representation. It can be thought of as a mental representation of a general type of system or situation and contains very conceptual information about the kinds of objects, people, etc. involved and the relationships between elements. Thus the ORO's schema might apply to the generic capability of the ORO and refer to the kinds of things typically found in an OR, the types of actions and processes, and the relationships between things that go on in an OR. The schema would NOT refer to any particular OR in the context of a particular process, and no real OR would have all the elements of the schema. Examples of schema that the OR may use include: concepts of operations and knowledge of specific processes (pre-planning, operational procedures).

The relationship between the 'mental model' and the 'mental picture' is usually uni-directional, that is, the mental model will often drive or constrain the nature of the mental picture. This may occur in several ways. The model may serve to guide selective *extraction* of required information from the operating environment, and the model may also direct the way in which the ORO creates new mental representations from existing data that is internally re-assembled or re-configured. If the ORO only has available uncertain data or impoverished mental pictures of the external environment, then the ORO's comprehension of the current situation will be necessarily degraded, even though the model itself may be valid. However, under some circumstances the mental picture may also constrain the mental model itself. For example, it has been shown in the nuclear industry that an operator's mental model that is based solely on the information derived from control room displays, is likely to be insufficient in finding solutions to unexpected operational problems. To comprehend the problem and generate appropriate solution options a deeper understanding of the basic nuclear generating process is required.

Finally, it should be noted that it is unlikely that a mental model cannot be generated from a single mental picture, because there is insufficient information; thus, a number of mental pictures are usually required to generate the mental model.

The relationship between the *model* and a *schema*, is that the mental model can be generated from a schema to create a model of a specific system based on what is known of that kind of system in general.

5.1.3. Mental representations in the context of the OR

In the present context, the mental representations of the ORO were described during the CTA in many different ways. In some cases, ORO's described a mental model that comprised a number of mental pictures with *iconic*-like properties arranged and organised according to the model structure. In other cases, the model appeared to be more complex and took the form of a *structural* representation of knowledge, processes or plans. For the experienced ORO this may range from
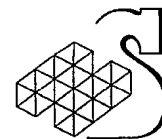
something as simple as an acronym that represents a list of things to do, to a complex combination of rules or hunches and/or constructed mental pictures of some part of the outside world such as the ocean floor and thermal layers which might conceal a subsurface contact.

These *structural* elements of a model may not only comprise a plan of action and the "list" of information to acquire and manipulate for a particular context, but also contain the relationships among items on the list. Within such a structure, the model incorporates the current information (data content) for each item on the list. This may be illustrated by a simple example. In the case of a new contact, the mental model may comprise: (i) a standard plan for action for the ORO, to ensure that the OR follows the appropriate pre-planned and trained procedures, steps; (ii) a "what-if" plan to anticipate future events and (iii) the data content comprising an up-to date mental picture of contact's course, speed and proximity to the TG and (iv) an understanding, and low level monitoring, of ongoing background tasks.

Of course, both structural and content aspects of the model will be subject to change. The structural aspects of the model are probably initially developed during training and evolve as result of experience. 'Structure' can be expected to be more stable than 'content' and to change more slowly once established. Individuals will discard what does not work for them, and add what does as they try different mental strategies to make sense of the information that they have to manage. 'Content' will change frequently, as new information is detected or required. Structure will not only determine how information is comprehended and inter-related but also guide search patterns to acquire and update content. The close relationship of these ideas with the acquisition and comprehension of Situation Awareness will be apparent.

For present purposes, the initial analysis in this report largely addresses the information" content" issues of such models. However, the overall schema that includes knowledge-representation structures and plans will be as important to understand, if the goal is to improve overall support for the ORO's cognition in future system design. The manner in which such mental representations are developed, what influences their development and degradation, the degree of homogeneity that is desirable or possible are all important design related questions. The degree to which any mental representation accurately represents the way the "real world" works will be critical, as in the nuclear generation example provided earlier, where the operators' model reflected a limited and insufficient view of the nuclear process that had been derived mostly from displayed system data. Without the answers, it will be difficult to achieve the best possible cognitive fit between OROs (or any other user of naval C2 information systems) and the presentation of the information which a person in that role is expected to manage. The term 'cognitive fit' is used to denote the degree to which the configuration or display of information matches the needs of the user as they gather and use it. A poor cognitive fit would be characterised by higher error rates, longer response times, more mental calculations or transformations on the part of the user to integrate elements of information, higher mental workloads, more items of information forgotten.

In the following section we attempt to provide examples and a simple categorisation of the way in which members of the OR mentally represent or integrate different types of information, with the caution noted that there is no single satisfactory taxonomy that meets all needs. The taxonomy is organised along the three dimensions of: '*level of abstraction*', '*level of focus*' and '*domain of interest*'. While the categorisation provided below may need to be refined, it allows us to address measurement issues more precisely for now. Note that these categories are unlikely to be orthogonal, i.e. they overlap and interact.
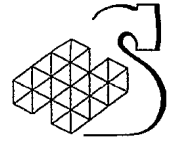
5.1.4. Level of abstraction

This aspect of the mental representation concerns the relationship between the representation and the data on which it is based and is reflected along a *concrete-abstract* dimension. For purposes of explanation we describe four broad categories of representation along this dimension.

In its most concrete form, OROs described a '*mental picture*' of the external world. Within this picture, several key elements of the environment may be visualised together with their spatial relationships to each other. An example of this would be the various positions of the vessels in the Task Group and their relationship to the maritime environment. This mental picture itself may take various forms. It could be a direct visual analogue of the plot, or the current symbology display on the CCS. A more complex picture might involve an integrated image assembled from data from a number of audio or text messages (such as the OOW) or display sources (threat board), embellished by, and integrated with the OROs last visual recollection of the outside world and previous intelligence reports.

As we have described earlier, in its most abstract form, the mental representation may be more of a schema, whereby the ORO constructs a mental plan of tasks to be done, their time frame, sequence, outputs, resource requirements and constraints. For illustration, we provide several examples below along this concrete/abstract dimension.

Category 1 – Most concrete
At one end of the concrete-abstract dimension OROs describe a mental picture that accurately images or reflects the data on which it is based. Such a representation could be an image of the plot, a situation display or a state-board etc. The important element of this type of representation is that the ORO can internally "read-off" the data from the picture some time after turning attention away from the original source. The data represented in the mental picture may map in almost a one-to-one manner to the data from which it was constructed, and allow the same degree of analysis by the ORO, as if the original source were still being viewed. An example of this type of representation would be a mental picture of the relationship between the ship and a navigational landmark. In this picture the bearing and distance information are spatially represented in an accurate manner, thereby allowing the ORO to make a navigational decision without the need to consult again the original source. It is possible also that the ORO may be able to recall this form of representation some time after acquiring the initial data, but with loss of some degree of detail. The number of intervening events that demand attention will probably degrade such images.

Category 2 – Concrete and abstract
Moving along the concrete-abstract dimension, we can consider a form of representation that also has picture like qualities but represents some degree of abstraction from, or integration of, the data on which it is based. An example of this might be a general picture of the task group as a single entity, its relationship to a distant surface contact that has launched a missile that is closing and the appearance of an unidentified air radar track. Such forms of mental picture might typically combine specific aspects of information across the maritime and air domains. In order to construct such an image the ORO would have to select and compile information from several data sources each with its own interface format (i.e. method of displaying the data). The ORO may also need to integrate relevant information that may have been acquired through auditory channels (for example, an auditory communication from the helo indicating distance and bearing to the target) with information visually acquired. This ensuing constructed image may represent spatial relationships between the entities in a manner that enables the ORO to comprehend a bigger picture; however, it lacks the precision of the icon-like image with respect to detailed information (such as exact distance and bearing). Such images will also be dependent, in large part, on the
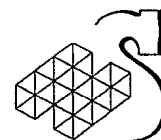
format and level of fragmentation/integration with which the information to be integrated is presented.

A further feature of images of this type is that to some degree they may be able to be manipulated to reflect different perspectives, both in spatial/temporal terms as well as tactical. Thus, the ORO can add data and manipulate the mental picture and project it forward over time, for example in envisaging a planned intercept course, or the way spatial relationships might change as a result of extrapolating speed and bearing.

No matter the source or the format of the information, the types of mental picture described above represent for the ORO an internal image of some aspect of the environment that needs to be retained over time and recalled from memory when required.

Category 3 – Mostly abstract

At a higher level of abstraction, the mental picture represents situations described by OROs in which they envisage aspects of the task environment in which conceptual or meaningful elements are more important than spatially veridical relationships (e.g. distance, exact location, bearing). Examples might include: visualising damage or potential damage to various areas of the ship, or task group, and projecting the consequences, memorising the status of equipment, or equipment characteristics such as radio frequencies. In this form of representation gross aspects of the environment are represented with only a sufficient level of detail to serve the goal at hand. The goals might be; for example, to establish which damaged area should get priority, or how to work around a non-functional piece of equipment.

Category 4 – Models and schema

These forms of representation are more difficult to describe in detail or to categorise simply. They are used by OROs to support a range of functions that involve planning or anticipating future actions or processes, and where the structural aspects of the model and conceptual understanding tend to be more important than the associated data. These models may support overall mission-related issues such as mission intent and planning, Rules of Engagement (ROE) and how these might evolve as circumstances change, and additional knowledge is acquired about the enemy, own force resources, and surrounding civil activity. Some of these mental models are quite stable over time and circumstances, such as the pre-plans for the ship and the TG and the situations in which these plans would apply. Some forms of these models may also be quite clear cut and represent more or less instantaneous and largely procedural plans for responses to, for example, a torpedo in the water or an incoming missile. Other examples of this type of model include the OROs' use of a mental "GANNT" chart to assign OR personnel resources over time and to prioritise the ORO's own tasks, or the development of a plan to guide sequential sub-tasks for a segment of the mission, or for a special procedure such as the ZIPPO procedures.

Other representations are more labile in structure and may be used by OROs to plan in ambiguous circumstances. Such models may be mission specific and be built up over time largely during planning and preparation. The model is updated with new data, for example, as intelligence arrives, mission intent is modified, ROEs change or the political situation fluctuates. The content and structure of the model may be supplemented by more general background insights based on, for example, familiarity with the way in which this particular CO, TG commander or TG warfare director tends to react; or previous knowledge of this type of threat or circumstances in this particular theatre.
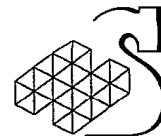
5.1.5. Level of Focus

The second, independent dimension of the mental representation that emerged from the CTA concerns the level of focus. The most typical example encountered was a distinction OROs frequently made along dimensions of scale, range or time. For example, *"ranged in"* represents a close range view within the maritime or air domain. This is usually associated with a "zoomed in" display and is used for a variety of purposes where the time frame is short or distances relatively small. The representation is frequently an internal, picture-like image that preserves essential object information and spatial relationships (Categories 1 or 2 from along the concrete-abstract dimension). The information may be a direct representation of the way data are presented on a single display, or may combine some selected elements from different sources. Examples where such an image would be created, and referred to, might include local navigation with respect to other ships in task group, other close surface vessels or navigational hazards and shoreline. In an air context it might be related to positioning the ship to minimise its radar profile and tactical manoeuvring for a missile threat.

In contrast, *"ranged out"* might also be called "wide range", or "zoomed out" view within maritime or air domains. This form of representation has all of the characteristics of the local view except that the scale (time or distance) is greater. It may be associated with a zoomed out display and is used for a variety of purposes, such as when the ORO is trying to build a picture of the relationship of new contacts to the ship, to task group and other selected environmental features.

5.1.6. Domain of interest

The third dimension along which mental representations differ concerns the warfare domain of interest. We mean by this whether the ORO's attention is directed to primarily air, surface, sub-surface or to some combination of these. At any one time an ORO may be trying to maintain a mental picture for each separate warfare area, within a model that provides an overall purposeful plan of action. Depending on the contextual circumstances, this picture may also vary along the degree of abstraction and also the level of detail. This within-warfare domain mental picture can be contrasted with what OROs frequently called the *"global picture"* or the *"big picture"*.

The **global picture** itself may have various forms of complexity and we provide two contrasting but complementary examples. At one level it is an *integrated* mental picture composed of a number of pieces of information selected from different displays or other data sources and integrated across surface, sub-surface or air warfare areas. Hence, it is relied upon extensively in multi-threat contexts as well as in situations where the overall mission intent and strategic picture may be the areas of focus. This form of representation is probably less detailed than the mental picture that the ORO has for any individual warfare area. The range of this representation could be close or far. An example of an integrated global picture might be a composite view of the relationships between the current position of a recce helo, a possible sub-surface contact and the task group, the potential weapons range of the contact and the endurance limit of the helo.

In addition to such a specific, threat-oriented global picture, OROs also described higher level global pictures. These tend to incorporate dimensions that are mission and resource related, take a task group rather than a ship perspective, and involve some appreciation of relative priorities and the overall global threat situation. This higher-order global view influences the formation and selection of lower level integrated pictures that deal with specific threat circumstances, the particular management of resources and operationally significant issues from a single ship perspective. Thus, the ORO may need to be continuously moving up and down this internal

---
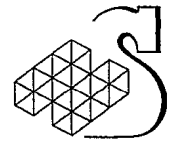
representation hierarchy in the achievement of local goals in the context of broader mission objectives and the needs of the person with whom he is interacting.

The ability to build global mental pictures of the types just described, is driven by the ORO's underlying global mental **model**. Overall, this represents a potentially broader comprehension and conceptualisation of relative levels of potential threat and interpretation of contacts. In the broader global context, the ORO may also need to call up other models, for example, to assist the global planning for the assignment of resources components and to plan the expected actions within each warfare domain.

### 5.1.7. Research and design issues

5.1.7.1. Development of mental representations

It is likely that much of the structure of the ORO's mental models is initiated during training and further evolves as result of experience. The manner in which mental representations are developed, what factors influence and shape their development, the degree of homogeneity that is desirable across individuals, or possible, are all important design related questions. Without the answers it will be difficult to achieve the best possible cognitive fit between OROs (or any other user of naval C2 systems) and the presentation of the information which a person in that role is expected to manage.

This perspective stems from several observations during the interviews in which we observed clear individual differences in ORO mental models even under similar circumstances. Such differences frequently appeared to result from different career paths to become an ORO (surface, sub-surface, navigator, communicator). Furthermore, as ORO's acquire experience in the role, there is ongoing development and changes to their meta-cognitive approaches to managing the information. Thus, their mental models will evolve accordingly. As individuals develop practical experience, they refine the way they think about things. They appear to be better at selecting key items of information on which to concentrate, to integrate them more effectively, to "see" things more quickly and simply, to hold important elements in the foreground of their attention, to be able to move up and down among levels of detail more effectively, and to be less likely to be overwhelmed with the schedule of things to be done. This description fits well with what is known about skill acquisition for other complex domains of knowledge.

Such differences will need to be accommodated through some combination of training and design. The designer needs to know what techniques of data fusion will make it facilitate the creation and maintenance of the range of mental models likely to be held by different users, and by a single user in different circumstances. Trainers need to understand how to develop the appropriate mental models given the opportunities provided by the designers and the backgrounds of the trainees. Individual ORO's need to understand how to select subsets of available data and to configure a particular display format to support the mental model required to deal with the situation. Hence, it is both a research question and a training and design issue of how best to support the initial information selection and model building process.

In the context of measurement, the above analysis suggest that a number issues need to be considered in the development of MOPs; these include identifying and "controlling" for ORO background during data capture or analysis and providing a broad range of contexts under which different mental models might drive different data selection strategies.
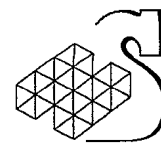
5.1.7.2. Selection of information content

The OROs reported during the interviews that superiors often demanded '*information rather than data*' from their support teams. It was also noted that a distinguishing feature between experienced personnel and novices, and between newly formed teams and experienced teams, was the tendency for the novices to report data rather than provide information. Explanation of the nature of '*information*' varied somewhat across interviewees, but some examples of '*information*' provided included: comprehension or insight into the relationship between data items, the projection of trends, making deductions etc. Generally one might say, therefore, that they described *information* as the comprehension of the raw data in the context of the present circumstances, or disparate data meaningfully integrated.

Given this distinction, C2 systems that display data, as opposed to providing information, afford little support for the establishment and ready implementation of a mental model. Thus, for the future, one key element to providing effective system support for mental models will be to identify the critical information needed by the ORO for the creation of the model for the mission related task in question. This information is difficult to describe generically since it will change depending upon a variety of circumstances. For example, in some cases the ORO may need to visualise the immediate spatial relationships between selected objects of interest (for example in planning a manoeuvre) such as the changing relationships between a fast moving contact, the ship and the task group. In other cases the ORO may need to build an internal array or matrix of abstract mission or resource related information.

5.1.7.3. Maintenance of mental representations

Little is known about how long an ORO needs to, or is able to, maintain any given mental image or model before it needs to be updated. One aspect of this is the *structure* or framework of the model on which current information items are hung and which guides the search for new information. Another aspect is the currency of the *information* that fills the structure and provides a mental picture. OROs described problems with loss of the mental picture under a variety of circumstances, particularly in situations involving high information rates or prolonged intervals of inactivity in that domain (see below for issues related to switching between pictures). Research is required to understand: how the picture degrades over time (particularly when the domain of interest, or level of focus, is unattended), how picture maintenance is influenced by the rate of new information, and what pro-active strategies can be adopted to maintain the picture. The impression left by the CTA is that the maintenance of the image aspects of the representation may pose a greater challenge to the ORO than the maintenance of the underlying model.

5.1.7.4. Switching among mental representations

Even when the operational environment is slow paced and the task focus narrow, the ORO must repeatedly switch among mental representations to select and attend to specific information related to different goals. This switch may involve a change in range, a change in domain of interest (air/surface), a switch between an ongoing background task (e.g. review of text messages) and a newly emerging situation, between a representation of the present situation and the projection forward over time (e.g. planning ahead), and between the strategic and tactical.

The need to re-establish a mental picture typically occurs after an ORO has been concentrating on one particular goal. For example, in watching the team deal with a pressing, unknown, but potentially hostile air contact, which engages in series of manoeuvres that bring it ever closer to a weapons threat range, the ORO may neglect the surface or sub-surface picture for a time. On
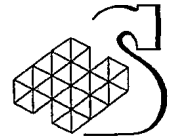
returning to the neglected domains, it is not clear how much previous information is retained or whether the picture has to be re-assembled from scratch. OROs consistently reported a "loss of the picture" in one area of interest when they have been concentrating on something else.

Overall, the CTA results suggest a number of research questions to be pursued:
- How frequently does switching occur?
- How long does the ORO stay with a single representation?
- Under what conditions is there interference or facilitation between pictures?
- What is the worst-case situation for frequency of switching?
- How do OROs update a pre-existing mental picture after switching back?
- Can switching (or its negative consequences) be reduced by improving cognitive fit with situation displays?

Any research that is performed will need to consider the complex reality of operational circumstances with moderate to high information rates. As the number of potential threats or contacts of interest increases, or the complexity of the tactical situation increases, the need to maintain multiple mental pictures and to switch among them also increases. OROs report that under such circumstances loss of all or any of the pictures is not only common but also stressful. Presumably, this stress is not only because of the high workload, but also because it involves loss of control in a high-risk situation. Such perceived loss of control has been shown to be an important negative factor for continued and effective problem solving under stress as well as a stressor in its own right. Such issues seem likely to be particularly important in understanding how OR teams deal with two or more simultaneous threats. Such circumstances were identified by interviewees as being probably overwhelming and impossible to manage with current systems. Given the likelihood that such tactics (i.e. attempting to overwhelm the information handling capabilities of the target organisation) would be actively pursued by any enemy, this issue seems to merit high priority for further investigation.

5.1.7.5. Scaling and spatial information

There are a number of issues related to how the ORO maintains, or needs, accurate information concerning distance scaling (or range), and/or time scales, within mental models involving spatial representation. The CTA data suggested a number of research and system design related issues. These examples represent an initial identification of potentially important issues, rather than a comprehensive list based upon detailed CTA probes or a review of the relevant literature.
- Under what circumstances does the ORO try to visualise and maintain scale information?
- Can the ORO accurately maintain a representation of scale within the picture?
- Is the accuracy limited to selected items within the picture or required for all elements?
- How does switching from a near to a long-range focus, or between air and surface domains interfere with the maintenance of scale information?
- Can the ORO acquire information represented by the system at one level of scale and mentally scale up or down the spatial relationships to the necessary level of accuracy?
- How are range and time scales inter-related and best represented? (For example, the need to co-ordinate changes in range scales with flight time to maintain screen observation of the track of a long range missile as it progresses along its incoming path.)

5.1.7.6. Projecting mental representations into the future

Under certain circumstances the ORO must transform the current representation to envisage some future state (i.e. Endsley's level 3 situation awareness). Simple examples include, envisaging
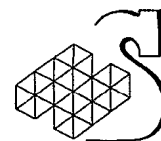
course changes to maintain position in the task group, while avoiding local hazards or obstacles, and projecting the plot of an intercept course to or from objects. A more complex example might involve planning an optimum recovery plan for a helo low on fuel while positioning the ship defensively to shield a high value asset from an approaching threat. This would involve taking into account the speed and movement of one's own ship, the high value asset, other ships in the task group, and possible hostile responses from the approaching threat with respect to all of the above.

A design issue is what kinds of tools would be needed to support this process of forward projection of the model over space and time.

5.1.7.7. Sharing and communicating mental representations

In many situations, several members of the OR may construct or want to share the content or structure of mental model for a particular area of interest- usually a picture of spatial relationships. One reported problem for the ORO is to ensure that the picture that the team is discussing is indeed "common" and does not differ in terms of scale, or layers of visible information. Two members of the OR team communicating about a common area of interest may have each of their displays configured slightly differently and hence create ambiguities not only about content but also location of key data. (There may also be analogous circumstances concerning non-spatial information.) OROs and others have reported physically moving over to one another's displays, or to the plot to establish a common picture. Frequent reference was made to a pre-existing technology that grouped key members of the OR around a common plot. This was perceived as having advantages in respect of such a common picture. (The old, common, plot also permitted face to face contact and more ready interpretation of the comprehension of others based on body language and facial expression).

Even when a common picture is established, OROs express some frustration in not being able to effectively communicate details among the members sharing the picture. From a design perspective, any future system needs to support the rapid sharing and annotation of images of common interest, but to achieve the appropriate design, a number of issues will need to be addressed through research. These issues include: how to transmit visual/spatial information effectively over audio channels (the need for a common short-hand language), the development of a tool suite to support creation and sharing of images, as well as basic research to understand important characteristics of shared mental pictures.

5.1.8. Summary

In this section we have outlined several critical components of mental representations which we believe to be a core, cognitive construct to emerge from the CTA. Mental representations are central to the ability of the ORO to manager and lead the OR and influence strongly the success with which these tasks are done. We describe in some detail the ways in which mental representations are used in an operational context and we provide a preliminary categorisation and definition to assist the process of developing MOPs and informing future system design issues. Finally, we identify a range of critical issues where further research and analysis will be required in order to provide data to guide future system design issues.
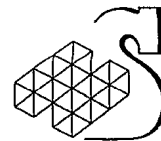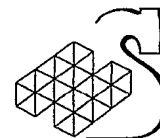
5.2. Data tables

The information acquired during the interviews was collated and reduced as described in Section 4.2.5. The outcome of this process is shown in the Annexes to the report. This section describes the general format of these Annexes.

An important background insight to the following is that ORO interviewees repeatedly emphasised that the role of the ORO was to maintain awareness of the "big picture" and to manage individual threat responses using OR team resources through the warfare directors rather than take over the response to individual threats. The ORO would only become involved in detailed threat response in the event that one or other warfare director was unable to cope satisfactorily. Furthermore, in the event of the ship coming to Action Stations, additional resources in the form of off-watch OR teams would arrive in the OR and come under the direction of the on-watch ORO. A feature of the scenario used, and one that distinguishes the littoral setting of this scenario from deep sea operations, is that the OR team would be dealing with data from a more or less continuos flow of surface and air contacts and underwater noise. The circumstance in which a single contact emerges on a more or less blank screen was not considered.

- **Coming on watch tasks (Annex A)**
  These goals / tasks represent the need of the ORO to update him or herself in all aspects of OR operation and awareness of the tactical and strategic situation for the mission. In terms of the foregoing discussion, the ORO is updating the content of previously established mental models from the point at which he or she was last on duty in the OR.

- **Background tasks (Annex B)**
  Background tasks are primarily those associated with the ORO's responsibility to manage the general preparedness of the OR to meet any threat and must be performed throughout the watch. (Foreground tasks, by contrast, are associated with meeting a particular threat using the capabilities of the OR.) In general terms, these tasks relate to several needs of the ORO such as the following. Monitor and manage the OR team. Manage information from all sources coming to the ORO position. Monitor and manage OR equipment capability. Achieve and maintain personal awareness of the strategic and tactical situation. This involves monitoring message traffic, OR activities, and the tactical picture (both directly through selected sensor data and indirectly through the reports of team leaders in the OR). Some background tasks (such as reviewing incoming text messages) typically make significant and continuous demands on the ORO's attention.

- **Foreground tasks (Annex C).**
  These goals / tasks start once a specific threat is detected and tend to be initiated once pre-determined criteria or trip points require the pertinent warfare director to report to the ORO.

Each of these Annex contains data tables following the format outlined below. Annex D relates the generic goals to specific event sequences taken from the scenario.

ORO Goals					
Criteria:					
Awareness needs	Decisions required	Communication sources	Mental Workload	Measure	Method
Comments					

Table 3: Outline Data Table

An introductory section at the start of each Annex outlines the general area of behaviour in question and some current contextual issues. For example, for reviewing incoming text messages, we attempt to provide some sense of context as provided by the interviewees in terms of traffic volume, frequency, message length, etc. Selected data tables contain supplementary notes. Within the tables themselves the following headings are used.

ORO goal(s): The specific objectives that the ORO expects to achieve.

Criteria: These provide benchmarks by which the goal-directed behaviour may be assessed. For the most part we have defined mission-descriptive criteria, such as performance quality and accuracy, goal accomplishment, time to perform tasks, response time etc.

Awareness needs: In this column we record the state of information awareness which the ORO is trying to reach as a first step in guiding the goal directed behaviour. Also recorded are the critical pieces of information needed to support situation awareness.

Decisions: The major and critical decisions are listed. Some decisions result in explicit actions and commands to other members of the team. Other decisions may guide the ORO's own subsequent behaviours. (e.g. to consult another screen or data source in order to provide more data for the task in hand, assess a threat).

Communication: The major modes of communication and sources/destinations of communications are listed. Communication is interpreted in the widest sense: namely to access information external to the ORO – whatever its source. This is usually expressed in terms of the current source(s) of the information required. In future systems, sources may change. For example, to obtain information about a radio frequency, the ORO might ask the ORS or a signaller, look at a state board, or check a post-it note pasted by the CCS screen. All such activities have been observed.

Mental Workload: Using ORO responses and our own judgement, we estimate whether the tasks at hand are likely to involve periods of high or unmanageable mental workload.

Measure: We provide a list of the appropriate performance indicators to assess the adequacy of the behaviours in achieving the required goals. Measures are related to those aspects of situation awareness, communication, workload and decision making which are core components of the task.

Method: Each of the methods appropriate to the measure in question is listed in summary format. See following section for further details.

Comments: In this column we provide a range of relevant comments. These might include: implicated cognitive processes; team management issues; expert/novice differences; particular operational problems reported; design issues and topics that may merit research. We also identify some possible areas where improved system design may enhance the decision-making needs of the ORO.
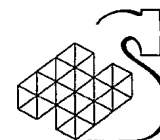
6. Measures and Methods

A preliminary outline of proposed measures and methods was provided in Reference B. The goal of this section is to provide additional detail on the proposed approach to collecting performance data and to provide an explanation of the information contained within the CTA data tables. The specific MOPs described are organised below in the four categories of situation awareness, decision making, communication and mental workload.

6.1. Criteria, Measures and Methods

Evaluation can be considered from several perspectives. For each ORO goal, the first perspective will be to establish a "Criterion", which identifies a broad dimension of evaluation interest. Next, to focus and operationalize the criterion, a specific "Measure" and "Method" should be devised or selected. Finally, a "Standard" will have to chosen by which to interpret the data.

In some instances the criterion may remain the same while the measure, method and therefore the standard vary. For example, to evaluate a criterion of "awareness for new information" for a mock up of proposed display format, a rating scale might be used by experienced SMEs as they walk through a standard scenario to estimate acceptability. The standard might be that the mean scale response (taken from at least four, independent, SME reviews) on a five point scale must be at least 3.5 (with 1 being "completely unacceptable" and 5 being "completely acceptable"). Later, using a working prototype for the display, the same criterion might be evaluated by measuring the time taken to respond to incoming information or accuracy of recall for information with the standard expressed, respectively, in seconds or the number and type of errors. Standards cannot be chosen at this stage and necessarily will involve SME operational insight. Moreover, Standards may change as weapon sensor systems develop and further response times or greater areas of surveillance become available.

CRITERIA	MEASURE	METHOD
Speed of awareness of new information	Time to show appropriate response to information change	User response time
Accuracy of perception of new information	% information misidentified	Track user mistakes and errors
Ease of use	% users rating feature(s) as "Acceptable"	Rating on scale by SMEs after performing standard tasks
Speed of message preparation	% users able to prepare message(s) within time limit for each workload condition.	Time users preparing standard messages under differing workloads

Table 4: Examples of Criteria, Methods and Measures

6.2. Situation Awareness

The analysis described below is focussed on ORO goals in the context of a specific threat. However, as noted above, the acquisition and maintenance of SA should be considered in a much broader context than a particular threat environment. For example, in the process of coming on watch the ORO updates SA for the upcoming watch by building on preparation, planning and mission experience. Further, in the ongoing operational environment, the ORO must continuously
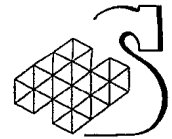
switch SA between background tasks (i.e. those that are required all the time – e.g. manage OR team, build global picture), and foreground tasks (i.e. those associated with a particular threat).

No matter what the specific context, the three basic *measurement* questions concerning SA are:

- does the ORO select and/or detect the appropriate information for the circumstances.
- can the ORO comprehend and integrate this information into the current mental model.
- when appropriate, can the ORO project the current mental model into the future?

6.2.1. Methods

Two forms of probe technique are proposed : a "*freeze*" probe or an "*embedded*" probe. Both methods have been used successfully and reported in the literature.

A freeze probe is derived from the SAGAT method (Reference D). Using this approach, a simulator scenario would be momentarily frozen (or an exercise interrupted) and questions posed to the ORO regarding current knowledge of factors deemed pertinent to all goals (foreground and /or background). The ORO may or may not, be allowed to interrogate the system to answer the questions, depending on whether one wants to assess current knowledge in isolation from the system, or the speed, accuracy and completeness with which information can be acquired. Such a probe may also end a scenario. Associated measures (completeness, accuracy, response time, etc) are discussed in the next section.

The advantage of this approach is that it directly queries the almost current state of knowledge, and virtually any aspect of the knowledge state can be probed. The structure of mental models may be inferred from this based on consistencies in the data from several different OROs and/or in conjunction with further diagnostic questions.

The major disadvantage is that the method disrupts ongoing information flow and risks inducing artificiality to the behaviour. As a result, anyone who know that they will receive such interruptions and probes may constrain and change their natural behaviours to meet the demands of the test environment. Further, once the interrupted mission segment is resumed, there is likely to be considerable overhead for the ORO to regain full situation awareness and resume tasks in hand. Therefore, this method poses some risk to the validity of the data obtained. However, there may be circumstances where this represents the only viable and valid approach to collecting the desired data. At the cost of increasing the number of scenario runs and the number of subject OROs to be tested, some of these disadvantages can be reduced by varying the predictability of the probe in terms of timing and content.

Variations on this approach are to tape the scenario and then conduct a cognitive walk through with the subject ORO afterwards and probe for their recollections of information used at predetermined points in the scenario. The disadvantage of this approach is that memory fades or responses may be distorted by information gained at other points during the scenario.

An embedded probe is a item of information or request for information injected naturalistically into the ongoing scenario by the measurement team that is disguised as part of the normal message traffic or flow of data. Data captured is the downstream response by the ORO, evidenced as part of his normal operational actions. In principle, there is no interruption to the normal state of activity within the OR. The embedded probe may contain relevant information about the operational environment that constrains and modifies future actions by the ORO. Alternately, it could be disguised as a request for some piece of relevant data or information from a contextually valid source. Depending on the exact nature of the probe and the behaviour anticipated, it should prove
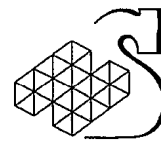
possible with strategic sampling of ORO responses to provide measures of all three components of situation awareness. The embedded probe method allows for a good range of flexibility both in terms of how and when the probe is delivered and the nature of the responses expected. A probe could have any of the following characteristics:

- information provided in a briefing package
- information provided in the watch hand-over
- a message from outside the OR
- a message from within the OR (by paper, direct voice or audio net)
- information embedded in normal displays (e.g. sensor displays, CCIS etc).

OR responses are collected further downstream from the injected data at anywhere from a few seconds to potentially hours delay. Many actions of the ORO could be a response to a probe and might include:

- direction to another OR team member
- the selection of a particular display screen
- a message within or outside the OR
- movement to another location
- a query to someone else in the OR team

Although it is not possible at this stage of the analysis to prescribe the circumstances in which each type of probe will be adopted, there is one major constraint that should be noted. In general, the longer and more complex the test scenario, the less control can be exercised by the measurement team over the evolving conditions. Hence, there may need to be artificially halt an evolving scenario at several points during its unfolding and to return to a baseline state. In such a circumstance, the freeze probe methodology would be widely used to assess the ORO's SA with each stoppage of the scenario. The method could be used simultaneously across an OR team to assess SA of several team members with respect to a common problem and to track the transfer of information among the team.

Another specific situation where freeze probes are required, is in a context where we may need to understand more fully the content of the ORO's situation awareness. For example, in a threat assessment context, in which no course of action is ultimately selected, it will be important to actively probe the information considered by the ORO and the decision alternatives reviewed. Freeze probes can also be used more widely for test contexts that involve shorter scenario segments, task clusters and task strings.

Note that this consideration entails an aspect of scenario/test environment design that will be done as part of future work, although it is assumed that simulation facilities for the system in question would provide an appropriate environment for data capture. We will also need to address important practical issues, such as when to use whole or part tasks, the complexity of the scenario segment and the scenario length. Different measurement approaches will also be required according to the state of system development. At the concept development stage, story boards or display mock-ups may be used for SMEs to conduct a cognitive walk through of a representative scenario while provide subjective comments or rough estimates of in response to probe questions. Later in development, partly functional prototypes may be used in conjunction with a partly simulated scenario to capture part task data. Finally, full sea or land based trainer trials will permit full measurement and evaluation of the entire system.
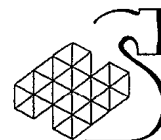
6.2.2. Measures

The potential *measures* of situation awareness include the accuracy, completeness, the latency and the order of reporting responses to the probes. Reports may require verbal, questionnaire, or in the case of spatial data, graphic answers. (For the latter, respondents might be required to sketch out data points on a chart or manoeuvre board.)

The following table illustrates selected criteria for assessing situation awareness, their associated measures and measurement method.

GENERAL FUNCTION	CRITERION	MEASURES	METHOD	NOTES
Maintain effective Situation Awareness				Can be broken down into the subsequent criteria
	Maximise detection of changes in status of key factors for awareness	Latency of detection	Response time to status change in selected variable.	Establish variables. Data capture using video or software.
		Accuracy of detection of changes in key variables.	Freeze or embedded probe @ selected variable.	
	Maximise comprehension of incoming information	Comprehension of information.	Freeze or embedded probe	Need expert SME to rate answers.
		Completeness of information provided to user	Complete check list for information provided by the system	Cognitive review of scenario by several SMEs
		Completeness of information acquired by user	Freeze or embedded probe	Embedded probes disguised as legitimate requests for information.
	Maximise ability to project situation into near future.	Accuracy of projection	Freeze or embedded probe	Need expert SME to rate answers. May be several acceptable answers.
		Latency of projection	Response time to probe.	Data capture using video or software.
		Completeness of projection	Freeze or embedded probe	Need expert SME to rate answers.

Table 5: Example performance measures for Situation Awareness

Other important criteria include the ability to switch focus of awareness and acquire/update/renew awareness for the new focus as quickly and completely as possible, minimise decay of awareness, and maximise team awareness in terms of completeness and commonality of detection and comprehension. The basic performance measure for most SA requirements is the degree to which accurate responses are obtained to probes about awareness for a list of information items deemed necessary for solution of a given problem. For example, potential probes for a situation where the ORO is directing the ship to protect a high value asset from an incoming threat might include:

- what is the intercept bearing between the threat and high value asset
- where are there any navigational hazards and what are they
- how long will it take to get in position etc.

For response time, we use the abbreviation RT in the table. It should be noted that this is not being used in the traditional sense of reaction time (i.e. measures of usually a few seconds or less). Instead we are interested in the total time taken to generate an answer to the probe in question.
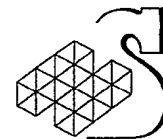
6.3. Decision Making

The measurement of DM presents a somewhat greater challenge than measurement in other areas. Some decisions may not result in observable behaviours (e.g. the decision not to look more closely (range in) on a contact of interest, or not to send a helo to investigate). Other decisions may involve complex assessments of a problem for which no single correct answer may be available rather a range of plausible answers. Some ORO decisions affect only the ORO's own behaviour, others result in OR actions that are expressed in command or control outcomes. Some decisions are related to closely antecedent events, others may occur minutes or hours downstream from the relevant circumstances (e.g. an enemy behaviour observed in a prior engagement influences the tactics chosen during a subsequent engagement; decisions made during planning are effected hours or days later during operations).

Given all of the above, there is no single measurement or method that can cover all of the possibilities, hence a variety of approaches must be used. No matter what approach is taken the primary issues of interest are decision quality and decision time. The measurement of decision time has greater importance for decisions made in a fast paced operational context, where slowness of action impacts upon system effectiveness. Such decisions may typically occur with high information rates and under the stress of threat conditions. Decision time is of less interest for decisions made during planning and preparation.

Decision quality is probably the more comprehensively useful approach to the assessment of the ability of the system to support ORO performance, since the ORO is the focal point of the many critical tactical decisions that are made during an operation.

6.3.1. Methods

For the measurement of decision time, the embedded probe methodologies described above is the procedure of choice. An embedded probe can be designed to trigger the need for a decision to be made that results in an observable outcome. The resulting time from the injection of the probe to the decision action can then be measured. The behavioural outcome that reflects the internal decision can take a wide variety of forms, that will need to be identified for the evaluation scenario. However, the following represent a sample of possibilities:
- the selection of information from a new source within the OR
- changing the display parameters of the CIS
- a verbal communication (see below for details on types) within or outside the OR
- moving to a different location in the OR

The next phase of work will be the pilot development of specific measurement protocols and methods. This will involve establishing a short specific scenario segment (or segments) to be run in some simulation environment. For the scenario segment chosen we expect to use SMEs (possibly instructors) to assist in establishing the appropriate behavioural measure for the decisions required within the scenario, and have each decision type triggered by a suitable embedded probe.

The measurement of decision quality cannot be as directly assessed from observable behaviours except for the simplest of decisions. For the critical and more complex decisions that are at the centre of the ORO's core role of directing the OR, suitable SMEs will be needed to assess decision quality. This should be systematised by asking SMEs to provide (in advance of testing) a listing of appropriate decision outcomes for the event. This would permit a measurement team to assess DM autonomously. However, some decisions may be only be assessed on a case by case basis by SME observers in real time, or by review of video records. In that case, to stabilise the data, more than
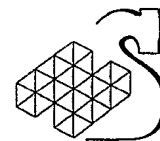
one SME (and preferably at least three) should rate the outcome(s) on standardised scales for each subject ORO.

In most cases it is assumed that decision quality will be assessed in terms of "satisfaction" of the problem in terms of the information available to the decision maker at the time rather than "optimisation" or in terms of the actual battle outcome (i.e. "winning" or "losing").

A variety of methods are available to implement the evaluation of decisions by SMEs and include:

- freeze techniques – where the immediate antecedent behaviour is probed and analysed
- real time unobtrusive observation and data recording by SMEs
- post event review of video, audio and data logs by SMEs
- post event debriefing of OROs by SMEs

The decision on which of the above methods to adopt for a specific evaluation scenario remains the focus of future work. Factors to be taken into account will include such things as logistical constraints (time, availability of personnel and test resources), the test environment and whether the situation involves, part or whole tasks, task strings, mission segments or extended missions.
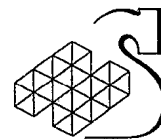
6.3.2. Measures

In the case of decision time, the RT to an embedded probe or to a known decision trigger is the measure of choice, providing that the decision outcome behaviour can be accurately detected and its time of occurrence recorded.

For decision quality, the assessment will in many cases be based upon an accuracy or completeness dimension. For example, did the ORO:

- Select appropriate information from among that available?
- Weight the selected information appropriately?
- Make a satisfactory choice in the circumstances?

For the latter, the "satisfactory" will need to be defined in dimensions appropriate to the scenario and the decision. For example, a decision to take over threat response from the SWC may be assessed in terms of a number of factors: (i) resource allocation for current and future demands on the OR, rather than mission intent and (ii) the need to defend the ship versus safeguard the high value unit being protected by the Task Group. In the meantime, in the CTA data table, we have simply summarised this accuracy component as "% *variable of interest*". The table below provides some examples to illustrate specific measures and methods.

| FUNCTION	CRITERION	MEASURES	METHOD	NOTES
Enable effective decision making				
	Minimise decision time	Speed of decision	Timings from insertion of critical cue(s)	Data capture: real time, video or software.
	Decision timeliness	Time ahead or behind schedule	Comparison to time line. SME rating for premature.	SME review: real time, video. Premature = did not wait for pertinent info even though time.
	Decision quality	Use of pertinent information	% pertinent info used	Checklist for specific scenario. SME review in some cases
		Information weighting	SME rating	Checklist for specific scenario. SME review in some cases
		Appropriateness of decision	SME rating	Checklist for specific scenario. SME review in some cases

Table 6: Example performance measures for Decision Making

6.4. Communication

Communication in the OR takes a wide variety of forms and involves a range of equipment. As outlined already, "communication" is viewed as any activity required of the ORO to access the information required for awareness in a specific situation. This is a broader view than normally taken but is necessary to permit comparison across systems that may employ different technologies to provide access to the information needed and to exercise control and direction. For example, what one system may provide over an audio net, another may provide digitally, and yet another on a state board. Measures of communication performance should permit comparison between such systems and relate to other measures such as situation awareness. Consequently no single approach is likely to capture the necessary range of communication activities. However, for present purposes, the illustration of measures and methods is largely restricted to auditory communication.

In considering the approach to be taken, the following factors need to be considered. First, no single communication mode is dominant. That is, information arrives and is transmitted regularly using all of the available modes. Second, auditory communication that goes through the ORO's headset may be on several channels, only one of which can be tape recording with the current technology. Third, much of the urgent messaging may accompanied by gesture or facial expression. Fourth, in many cases it may be impossible to ascertain with whom the ORO is communicating through video or direct observation. Fifth, there is no central log is maintained of all of the ORO communication.

Given the above constraints, the broad approach will involve capturing the basic elements of the more important aspects of the communication flow for the evaluation scenario in question. This will include when a communication starts, where it is directed, when and if it is received and how long it has taken, when and if it is attended to, and how well it is comprehended. Other aspects might include the degree of attention that must be invested in the communication process, and therefore is not available for other uses.
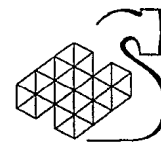
6.4.1. Methods

Video and audio taping of selected tasks and team members with subsequent review and data capture will be appropriate for many measures. Audio logs available from the training simulator may also be useful to track some selected aspects of audio communication. Such an approach will provide some basic performance measures relating to the flow and timing of critical information. However, this approach will be insufficient to address many of the specific questions concerning communication quality. Again, the use of SMEs to review behaviour in real time, or logged data, will be necessary. The use of embedded probes will be widely adopted to collect a broad range of data in a variety of communication contexts. SMEs will be consulted concerning the construction of the probe content and the appropriate time for it to be inserted into the information flow.

Since it will be impossible to capture all aspects of the communication flow during a mission scenario, we will continue in the next phases of the CTA analysis to narrow down the choice of which situations the measurement of communication should take priority and which measurement methods will be most appropriate for each situation.

As an initial effort to quantify communication activity, the intention is to classify fundamental communication patterns based on studies drawn from the Air Traffic Control and C2 literature (see table below). Although these categories were, for the most part, developed from studies of voice communication, with some modification they may be transposed for use in the broader context of communication outlined above.

Category A:	Command: Action required statement (e.g. move unit x to location y). Usually linked to "acknowledgement
Category B:	Inquiries: Queries, requests for confirmation or verification, other statements to elicit information Usually linked to "response".
Category C:	Acknowledgement: Affirms receipt of a command or inquiry (usually binary, e.g. yes/no, roger)
	Response: Offers or transmits information following a command or inquiry (e.g. ETA is 5 minutes). Conveys more than one bit of information
Category D:	Statement of intent: Communicates the expected or anticipated status of the mission. (e.g. Ready to go in 5 minutes.)
	Observations: Statements about obvious realities of the situation (e.g. target moving East at 15 knots)
	Suggestions: Advice that does not require mandatory action
Category E:	Non-task related: Joking, personal comments, social activities
	Not code-able:

Table 7: Categories of Communication

6.4.2. Measures

Based on this categorisation, several types of measure may be employed. Time-related measures of communication include the following.

- Percentage of time spent in each communication type
- The means and variability for transmissions of each type
  (this allows the identification of communication bottlenecks)
- Time between message reception and acknowledgement
- Time to respond to a query message
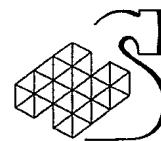
- Time between command message reception and implementing a required action
- Timeliness in sending messages to others requiring time sensitive information

To this may be added measures that are mode specific but probably impractical, such as the number key strokes required to access information in a computer data-base, or the number of eye or head movements required to capture visual information around the OR.

Errors and redundancies in communication can also be measured. For example:
- requests for clarification.
- repeated messages.
- incorrect read-back of a message recipient of a command, which goes uncorrected.
- read-back of a communication intended for a different recipient.
- command instructions given to wrong recipient.

Errors in message content can be assessed through reference to ground truth or by SME evaluation of video and audio logs of message activity.

In the case of the ORO who, in conjunction with the ORS, plays a central role in managing and co-ordinating information flow in the OR, the following areas of communication performance will need to be measured.
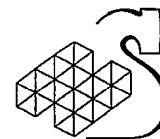- Prioritising communication
- Filtering communication
- Monitoring team communication

Based upon the comments received during the CTA interviews, it will be particularly important to take measures under high information rates or under circumstances which simulate equipment malfunctions. These are circumstances under which OROs may have particular difficulty in maintaining the necessary internal and external communication and are current sources of stress.

The following table illustrates some possible criteria for the measurement of communication, their associated measures and measurement method(s). However for present purposes, priority is being given the three measures that are expected to provide the greatest return. These measures will need to be related to the specific awareness needs of the ORO for the point in question of the evaluation scenario. These three measures are, to access to each item of information required:
- The number of communications.
- The duration of communications.
- The categories of communication.

Other communication measures may be added as appropriate.

FUNCTION	CRITERION	MEASURES	METHOD	NOTES
Maintain effective communication				Can be broken down into the subsequent criteria
	Minimise time spent in communication activities	Percent of total time spent on communications activities		Video/audio log
		Mean number of transmissions per system function	Video/audio log	May provide more detailed breakdown to diagnose communication problem areas
	Minimise attention required for communication activities	Time away from primary task Deterioration in primary task.	Embedded probe to initiate communication	SME review video record. Ability to communicate in parallel with other tasks.
	Minimise requests for clarification	Number clarification requests per "n" communications	Video/audio log	
	Minimise communication related activities.	Number of actions required to complete communication	Embedded probe to initiate communication	Depends on mode, key strokes, eye movements, length of verbal communication,
	Minimise the number of repeat messages	Number of repeat communications per "n" communications	Video/audio log	
	Minimise the delay in responses to inquiries	Mean time from end of inquiry to start of response	Video/audio log	
	Minimise communication generation time	Mean time from start to end of message composition	Video/audio log	Focus on standard reports e.g. after action
	Minimise errors in communication content	Number of errors per "n" communications	Video/audio log;	Compare to ground truth

Table 8: Example measures of Communication

6.5. Mental Workload

We propose to allocate lower priority to the development of workload measures and methods, compared with the other measurement domains. We believe that workload is less diagnostic for identifying system problems or to guide future system development (compared with looking directly at the processes that result in workload). On the basis of the visit to the simulation trainer and the CTA interviews, we believe that standard workload rating measures taken at key points during the evaluation scenario or a post hoc review with subject OROs will be suitable for the anticipated test environment. These measures and methods have already been outlined in Reference B.

We believe that the best initial use of workload assessment will be for the development of baseline performance statistics, rather than for diagnostic purposes. The early generation of such data for critical tasks will be important for later documentation and assessment of progress towards assessing the ability of future systems to optimise the ORO's workload. We anticipate that workload measurement may become a higher priority as the future system evolves and component functionality is brought together. At that point it may become appropriate to assess workload in a complex, dynamic environment that requires the ORO to multi-task effectively under moderate to high information loads and time pressures.
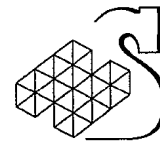
7. Discussion and Recommendations

This report outlines a framework for evaluating the ability of information systems to support the cognitive aspects of work conducted by the ORO. The analysis is based upon a scenario involving threat responses during the passage of an all-Canadian TG through a busy littoral environment. The framework is built around evaluation of three interactive processes underlying effective command and control (Situation Awareness, Communication, and Decision Making) with the Mental Workload of the user as a partial indicator of the effectiveness of the support provided for these processes. This framework for evaluation is interpreted in the context of the ORO's cognitive needs for fulfillment of mission related goals as related to background and foreground tasks and the management and conduct of the OR. ORO's accomplish this multitask management through the widespread use of mental models of the task environment. The report outlines some of the important, relevant information needed support the mental model in terms of SA, CM and DM and hence provides a direction for how to measure and evaluate these processes

Methods of data capture are based on what should be feasible in a dedicated test bed or training simulation context such as that provided by the Halifax land trainer, with some additional data capture capabilities. The proposed measures comprise setting up part-task scenarios with known information elements, about which participants are queried (directly by question or indirectly by reaction) at key points in the scenario to determine the state of their SA. This would be complemented by recording the communication patterns and content of participants and evaluated by determining the amount of time spent in communication, the number of communications, type, and errors in content. Finally, the quality of decisions made by would be evaluated by observers (real time or video review) according to a scenario specific checklist. Mental workload would be monitored at appropriate points. Information required for different ORO goals and associated measures are outlined in the Annex to this report, though details specific for any given scenario cannot be determined until the scenario is provided.
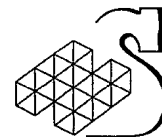
7.1. Mental Representation

7.1.1. Background in the Literature

Various forms of mental representation, mental pictures, mental models, and mental schemata, have been discussed in this report. Each has a rich research literature that has potential application to the development of the next generation of C2 support system for the HALIFAX class. To illustrate this, we have used the concept of the mental model based on preliminary research of the relevant literature.

The concept of the mental model has a long history in the psychological sciences and has been a topic of fierce debate (see Johnson-Laird, 1983). This concept arose as a theoretical perspective on mental representation and reasoning to rival formal logic and propositional theories. The advantages of mental models are that they accommodate analog or "fuzzy" mental processes that seem to account for human thought. Moreover, mental models are based more on an understanding of linguistic and general knowledge structures rather than formal rules, making them consistent with a wider range of psychological phenomena.

The development of the theory of mental models in psychology can be summarized in two ways. The first concerns human reasoning and contrasts the view that people employ abstract, general

rules of logic against the view that people reason with very specific, concrete representations of the problem domain (e.g., Johnson-Laird, 1997). Formal logical theories, of course, adopt the former view whereas mental model theories argue for the latter, considering all reasoning to take place in the context of specific instances.

The second approach concerns the nature of knowledge representation, especially the representation of discourse and text (e.g., Glenberg, Meyer, & Lindem, 1987). A traditional view is that people create propositional representations that mirror the underlying structure of language itself. The contrasting mental model view is that people create symbolic representations based on the *content* of language rather than the *structure* of language. The former view takes advantage of the appealing order of language, which can be decomposed and understood in terms of formal rules. The mental model view, however, bridges more directly the gap between language and reality, creating an internal symbolic system (i.e., a mental model) that simulates the systems and events described in language. (Although mental models are frequently studied in the context of language, they can be constructed from any kind of input, including direct perception of events).

The positions adopted by mental model theories on these reasoning and representation debates highlight the main features of mental models. They are mental representations aimed at mentally simulating the world. Thus the term 'model' is not just a convenience but an integral descriptor – mental models serve many of the same functions as physical models and have many of the same functional properties.
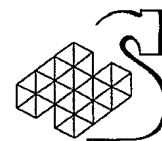
7.1.2. Research Issues

Despite over 25 years of active study within psychology, there remain many unresolved issues concerning mental models. Most notable among these is the question of how and when people use mental models. Although it is clear, as for the ORO, that people do use them, the more challenging questions of the situations and contexts in which people use mental models remain open today. A number of empirical studies have uncovered factors that limit the construction and/or use of mental models (Franklin, Tversky, & Coon, 1992; Gray Wilson et al., 1993).

Even in cases where the use of mental models is clear, we may not fully understand how knowledge is stored and organized. From the great flexibility of mental models comes some degree of confusion regarding how the mind actually takes and represents information. Further research is needed to illuminate the computational steps underlying mental models.

Similarly, it is not always clear how people use mental models and how they access stored knowledge. Studies of foregrounding (retrieval of information for working memory), for example, consider the relation of knowledge stored in long-term memory with working memory and the decay of that knowledge as a consequence of intervening information and events (e.g., Glenberg et al., 1987). The degree to which an association is created between two items of information in a person's mind seems to be an important factor. The related design issue is how this is achieved in a particular context for particular items of information.

7.1.3. Methods for Studying Mental Models

Studies of mental models have traditionally been aimed at addressing very basic theoretical issues (e.g., whether the way people represent knowledge is consistent with mental models or a propositional representation, etc.). However, while the study of mental models tends to be focused on very specific aspects of mental models, the methods used have potential for studying application issues, such as the needs of the ORO for a given context.

Perhaps the biggest issue in adapting empirical methods to the naval domain will be determining exactly what we need to learn about mental models in this context. Research in psychology focuses on basic issues of the existence and structure of mental models and these are both of practical interest to the design of OR decision support. For any given situation in the OR, we can ask whether people use mental models, propositional representations, mental pictures, or some other form of representation. Indeed, we can go further and ask whether people use the most suitable form of representation, whether their representation depends on the system or display, and how training can enhance (or impair) their ability to understand the situation. All of these questions have design implications; if we can find the kind of representation used, we can design systems, training team structures, and so on to accommodate people's representation strategies.

It is important to make use of the large psychological literature on mental models and related topics when addressing these questions. A great deal of research has explored these questions, albeit in more general contexts than that of the OR, and derived extensive empirical data. We need to find existing studies that have answered some of these questions. We have little doubt that many such studies exist but effort needs to be devoted to finding them and interpreting their results to the naval tactical domain.

Although many techniques address the basic issue of verifying aspects of mental model theory, other techniques have been designed to uncover the organisation of information in mental models. In this area, research in related fields, such as human factors, has been very helpful. Endsley (1995), for example, has developed methods to monitor situation awareness of air crews. Situation awareness involves a number of concepts not normally associated directly with mental models (e.g., detection of information items, stress) but is itself highly dependent on the concept of the mental model for comprehension and mental projection of the situation. Thus, techniques to assess situation awareness offer excellent opportunities to explore the structure of mental models in a wide range of settings, such as the OR. The reverse, of course, is also true as techniques to measure mental models can aid the study of situation awareness.

References for above section:

Johnson-Laird, P. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Johnson-Laird, P. (1997). Rationality, rules, and models. *Cashiers de Psychologie Cognitive, 16*, 114-123.

Endsley, M.R. (1995). Measurement of Situation Awareness in Dynamic Systems. *Human Factors, 37*, 65-84.

Franklin, N., Tversky, B., & Coon, V. (1992). Switching points of view in spatial mental models acquired from text. *Memory & Cognition, 20*, 507-518.

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language, 26*, 69-83.

Gray Wilson, S., Rinck, M., McNamara, T. P., Bower, G. H., & Morrow, D. G. (1993). Mental models and narrative comprehension: Some qualifications. *Journal of Memory and Language, 32*, 141-154.
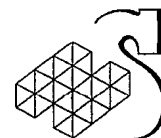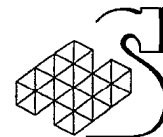
7.2. General Conclusions

The CTA provided a significant insight into the cognitive constructs which support the ORO's attainment of mission related goals. Some of these insights included the following:.

- OROs employ a variety of mental pictures and models to achieve their cognitive goals. Taken together, these mental representations vary in terms of level of detail, domain, and abstraction and may be modulated by career path, experience, other team members, mission planning, watch hand-over, and recent events. Better understanding of these mental models should underlie design efforts to achieve better cognitive fit with for the functional needs of OR personnel.
- Mission preparation has great significance for establishing mission related mental models. Support systems for mission preparation should be linked with support systems for mission implementation in terms of underlying issues relating to building and maintenance of users mental models.
- Updating Situation Awareness when coming on watch is critical for updating the ORO's mental models.
- Information systems need to provide less data and more "information" i.e. pertinent data meaningfully integrated in terms of the users mental model(s). This implies a need for functional support that can be adapted by "educated" users to individual, contextual and mission related needs.
- For Situation Awareness, major functional requirements (currently not well supported) include information acquisition and integration, cognitive fit of the available data to the mental model(s) of the user, regaining awareness after switches in attention between different areas of focus, and alerts for significant changes in unattended areas.
- Background tasks such as maintaining situation awareness of the evolving operation, or dealing with incoming text messages, can be differentiated from foreground or threat related tasks. These two categories interact and require the ORO frequently to switch attention and change mental models. OROs need to be able to switch between foreground and background tasks with seamless integration of data
- A major ORO function is to manage the overall OR team threat response rather than to be directly involved in the details of responding to particular threats (although this requirement may change for multiple threats if the ORO is required to assume the role of SWC).
- Common implicit intent and understanding among OR team members has particular significance for communication effectiveness.
- Handling of multiple threats particularly in areas of high traffic density is likely to remain a difficult cognitive challenge that may increase if the OR is assigned TG responsibility for a warfare area.

Many of these points probably apply, to some degree, to many OR positions.

7.3. Next steps

The next recommended step will be to verify the present findings and proposed measures with one or more experienced naval SME's. This would be followed by a brief, practical, trial of the measures and methods outlined in this report, using a small, selected segment of the present scenario. There would be three purposes for this trial. First, as a general verification of the measurement approach and to further refine the measures. Second, to check that data can be

captured using simulation resources available, first in the Halifax land based trainer and then, perhaps, on board ship (in harbour or at sea). Third, to work through the analysis of a mock data set to illustrate how baseline data might be derived for the present system.

A visit to the Halifax land based trainer suggests that this is a suitable site for the purposes outlined above. The trial would involve a brief review of this report, data tables and proposed measures with one or more SME instructors from the trainer to verify the findings. Based on this and still working with the instructor SME, a short representative scenario would be developed based on the present scenario and the specifics determined for each category of measurement: situation awareness, communication, and decision making. This scenario would then be set up on the training simulator and a walk through review of the scenario and its measures conducted to demonstrate feasibility of data capture and to capture mock data for consideration. Depending on the outcome of this trial, further investigations would be planned.

The literature on human reasoning appears to offer rich insights into the challenges faced by OR members and yet to remain largely untapped by naval decision making research. Therefore as a parallel step we recommend a review of the psychological literature on mental models and other forms of mental representations in the context of the management of advanced multi-task systems. Areas of this literature have evolved somewhat in isolation one from the other and from application domains. This review would contribute to issues related to both evaluation methodology, as well as research and design implications in the context of the OR.

#510918